

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2002, 07:44:50 ; Search time 1618.92 Seconds

(without alignments)  
16110.766 Million cell updates/sec

Title: US-09-603-124B-1

Sequence: 1 gcaggaacgcctccacggc.....gacgtaagtaacacagcag 1581

Scoring table:

Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
GenEmbl.\*  
1: gb.ba.\*  
2: gb.hg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.om.\*  
20: em.or.\*  
21: em.ov.\*  
22: em.pat.\*  
23: em.ph.\*  
24: em.pl.\*  
25: em.ro.\*  
26: em.sts.\*  
27: em.sy.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.hgo.hum.\*  
31: em.hgo.in.\*  
32: em.hgo.ro.\*  
33: em.hg.hum.\*  
34: em.hg.in.\*  
35: em.hg.ro.\*  
36: em.hg.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	1581	100.0	349980	6 AX127150	AX127150 Sequence
C 2	1558	98.5	2291	1 ABO15023	ABO15023 Coryneb
C 3	1545.6	97.8	4622	1 BLA242646	BLA242646 Brevibact
C 4	1458	92.2	1458	6 AX122446	AX122446 Sequence
C 5	1445.2	91.4	1458	6 E28467	E28467 Novel murc
C 6	392.2	24.8	20400	1 AEO07068	AEO07068 Corynebacteri
C 7	392.2	24.8	37586	1 MTCY270	MTCY270
C 8	392.2	24.8	37586	6 AX191745	AX191745
C 9	329.2	20.8	348450	1 MLEPRN4	MLEPRN4
C 10	329.2	19.5	38793	1 ABO03132	ABO03132 Coryneb
C 11	289.8	18.3	4116	1 BLFRTS2	BLFRTS2
C 12	280.8	17.8	5546	1 SC9B1	SC9B1
C 13	231.8	14.7	24800	1 MLCB268	MLCB268
C 14	226.2	14.3	38859	1 AP002997	AP002997 Mesothric
C 15	163.6	10.3	329709	1 SME591789	SME591789
C 16	162	10.2	294800	1 AEO08126	AEO08126 Agrobacte
C 17	141.4	8.9	14545	1 NMA72491	NMA72491
C 18	133.8	8.5	195767	2 LMFLCHR36-27	LMFLCHR36-27
C 19	132.8	8.4	110000	2 AEO02398	AEO02398 Neisseria
C 20	132.4	8.4	10629	6 AX044029	AX044029 Sequence
C 21	132.4	8.4	349980	6 AEO03919	AEO03919 Xylella f
C 22	130.8	8.3	20062	1 AEO03919	AEO03919 Xylella f
C 23	130	8.2	1476	6 AX189061	AX189061 Sequence
C 24	130	8.2	1998	1 ECUS7892	ECUS7892
C 25	130	8.2	2799	1 ECHMRGC	ECHMRGC
C 26	130	8.2	2157	1 AEO00118	AEO00118
C 27	130	8.2	28277	1 EC2MIN	EC2MIN
C 28	130	8.2	28277	6 AX191720	AX191720 Sequence
C 29	129.4	8.2	111401	1 ECOL110K	ECOL110K
C 30	129.4	8.2	10487	1 AEO05922	AEO05922 Caulobact
C 31	128.4	8.1	281530	1 APO02550	AP002550 Escherich
C 32	128	8.1	12518	1 AEO05186	AEO05186 Escherich
C 33	126.6	8.0	1550	1 AF306766	AF306766 Pseudom
C 34	126.6	8.0	1700	1 AE110740	AE110740 Pseudom
C 35	126.6	8.0	24000	1 AEO04656	AEO04656 Pseudom
C 36	119.4	7.6	22201	1 AEO04310	AEO04310 Vibrio ch
C 37	117.4	7.4	38510	1 AF179611	AF179611 Zyomonas
C 38	112.8	7.1	12085	1 U32794	U32794 Haemophilus
C 39	106.6	6.7	11584	1 AEO06049	AEO06049 Pasteurel
C 40	104.2	6.6	15862	1 AEO00736	AEO00736 Aquifex a
C 41	96.2	6.1	19619	6 AX067439	AX067439 Sequence
C 42	94	5.9	12394	1 AEO01920	AEO01920 Deinococc
C 43	92.4	5.8	312430	1 RPXX02	RPXX02
C 44	84.6	5.4	1713	1 ECHMRGC	ECHMRGC
C 45	84.4	5.3	21000	1 ABO52554	ABO52554 Shewanell

## ALIGNMENTS

RESULT 1  
AX127150/c  
LOCUS AX127150 349980 bp DNA 11-MAY-2001  
DEFINITION Sequence 7066 from Patent EP1108790.  
ACCESSION AX127150 AX114121  
VERSION AX127150.1 GI:114041138  
KEYWORDS  
SOURCE  
ORGANISM Corynebacterium glutamicum.  
Corynebacterium glutamicum  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;  
Corynebacterium.  
REFERENCE 1 (bases 1 to 349980)  
AUTHORS Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,  
Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.  
TITLE Novel polynucleotides  
JOURNAL Patent: EP 1108790-A 7066 20-JUN-2001;  
KYOMA HAKKO KOGYO CO., LTD. (JP)  
FEATURES  
source 1. 349980  
location/Qualifiers





QY 1568 agtgaacaagccag 1581  
 Db 1561 AGTGAACAAGCAG 1574

RESULT 3  
 LOCUS BLA242646  
 DEFINITION Brevibacterium lactofermentum murD (partial), ftsW, murG and murC genes, located upstream of ftsQ-ftsZ.  
 ACCESSION AJ242646  
 VERSION AJ242646.1 GI:6723455  
 KEYWORDS ftsW gene; ftsW protein; murC gene; MurC protein; murD gene; MurD protein; murG gene; MurG protein.  
 SOURCE Corynebacterium glutamicum.  
 ORGANISM Corynebacterium glutamicum.  
 Bacteria; Firmicutes; Actinobacteria; Actinobacteriales; Corynebacteriales; Corynebacteriaceae; Corynebacterium.  
 1 (bases 1 to 4622)  
 Ramos, A., Honrubia, P. and Gil, J.A.  
 Characterization and chromosomal organization of the murD-murC region of Brevibacterium lactofermentum ATCC 13869  
 Unpublished  
 2 (bases 1 to 4622)  
 Gil, J.A.  
 Direct Submission  
 Submitted (17-MAY-1999) Gil J.A., Microbiology, University of Leon, Campus de Vegazana s/n, 24071 Leon, SPAIN  
 Location/Qualifiers  
 1..4622  
 /organism="Corynebacterium glutamicum"  
 /strain="ATCC 13869"  
 /db\_xref="taxon:11718"  
 1..333  
 /gene="murD"  
 /feature="murD"  
 /function="peptidoglycan biosynthesis"  
 /codon\_start=1  
 /transl\_table=11  
 /evidence=experimental  
 /product="MurD protein"  
 /protein\_id="CAB66322.1"  
 /db\_xref="GI:6723455"  
 /translation="LKGADIAFLVKKHEQRIKALVIGADRAETVAALKENAPQASVFVTDKTEPFEMERIVTEAFSISEPDVTLAPAAASLDMEKMGORGLFAHNIITIKGLTEERG"  
 336..1808  
 /gene="ftsW"  
 336..1808  
 /gene="ftsW"  
 /feature="ftsW"  
 /function="cell division"  
 /codon\_start=1  
 /transl\_table=11  
 /evidence=experimental  
 /product="FtsW protein"  
 /protein\_id="CAB66323.1"  
 /db\_xref="GI:6723457"  
 /translation="MTGSKKKPARPNTGKRTGIGIRERISGANLLAPLTDVIMILCIYVILSCIGVWVYSSMTSIRBGSVNGTCVAGCINHYGFPAWVALLTRPOTIRNLNLIITVSYILLAVOIGICIGKEBEGSSWIALGIPQPSLAKVAIVMGAYLACGPGVQPHFNHNLRFQGVGFNAFLIFEGDGMASFLVLFMLFEGIMKRTIACVLLIALAVLAWGGGFFSSRFVYFALFGNHDVGIAPQSYOGFSLADSGGVLGLOSRAKWFYLPKAKNDFIALLIEELGLMGALVIALFACILYFGLRTAKSLDPFLGLMAATLTAASVQAFTNIGVGLPTVGIOLPMTISAGTSALITLTAAGMLISCAHREPEVSAASVQPAIDRLGLRPSSTLTSSNLSNMTKAAKXKSESSEVSGPLWACDRTPSAGATKMASWSTISSATLDM"

gene  
 1994..3112  
 /gene="murG"  
 1994..3112  
 /gene="murG"  
 /gene="murG"

/function="peptidoglycan biosynthesis"  
 /codon\_start=1  
 /transl\_table=11  
 /product="MurC protein"  
 /protein\_id="CAB66324.1"  
 /db\_xref="GI:6723458"  
 /translation="MKSGLRIRKANSPPKPRVAVAGGTAGHIEPALVAEALRDKH  
 GAVSALAGPELHLETPVPRKPNNDLKLPEPVAALGOARRA  
 LKPDQAAVIGGVYSRPAYMAKSLGPEFVEANRAGANKLGYKIGVGLNAV  
 AGSGMDVVGIPIRAVLSGARDASADRADRWGIDKRORTIPVTGSGSSTINKA  
 VEOADIDVACFOYLHVKKNELPAKRGYIPVPIIDDKMAITVADLYCSGRH  
 TVAEVYRAAQPALVYPLPRNGEDALNQAIVTACARQIDADPDLIDATDIL  
 LHPSTHSDSAKSTAGNASTVIAOMIATINSON"  
 3147..4607  
 /gene="murC"  
 3147..4607  
 /gene="murC"  
 /function="peptidoglycan biosynthesis"  
 /codon\_start=1  
 /transl\_table=11  
 /product="MurC protein"  
 /protein\_id="CAB66325.1"  
 /db\_xref="GI:6723459"  
 /translation="MTPHLDASODIDLSRVHLIGIGAGMSGVARILLANGKTVTGS  
 DADSRLLPLRNVGATVAVGHAENLELSGELTPVYVSPFAIPQDNPVARRBEG  
 IPVIRRSDLGLLEGSTOVLIASTHKTSTMSVYAMAAGNDSPAIQGNKAG  
 TNAHGTGEVFAIEADESDASLIRKENVVAVNVEPDHDFEKTPEAYFOVDFG  
 RITPNKILVGLVDAELGERSVKRGITVVGGRDVOVQHPVPMATVNSOYV  
 AECTRATINIDQAEVSVITQIPGDHMLNQAALLAGVYGVGDVNDKIVEGSDSGYR  
 RREFPGATIEGGKFMGAIVDDYVNHPTPEVTLRCAATRYAAKGVITATPPLIX  
 SRMEFOKEFAELSLADAAYLEITVAREQPDVGSSEITIDANTIVVIEPNSAV  
 PERIATAPNDIVITMGASVTHLAPEILIDOLNN"

BASE COUNT 969 a 1254 c 1317 g 1062 t  
 ORIGIN

Query Match 97.8%; Score 1545.6; DB 1; Length 4622;  
 Best Local Similarity 98.8%; Pred. NO. 0;  
 Matches 1557; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 ggaagtaacgctcacaagtgatgacagacatgattgccaatcaatagcaaac 60  
 Db 3047 GAGGTAACGGCTCCACGAGTGATGACACATGATGCTGCACATACCAAC 3106  
 QY 61 aactaaacgacagcctcaacgaaagaaatagttaaagtgaccacccacacttga 120  
 Db 3107 AACTAAACGACGCTCAACGCAAGAAAGAAATAGTTAAAGTGACCACTGCA 3166  
 QY 121 ttctggcaagaatattgattcttcccggtccactatggtattggcgagcggaat 180  
 Db 3167 TTCTGCACAGATATTGATCTGCTCCGCGTCCACTCATCGTATGGCGGACCGGAAT 3226  
 QY 181 gtctggcgttggccgaatctcttgcgcgctgaagaagcaagtcagttccagatcca 240  
 Db 3227 GTCTGCGGTGGCCGAATCCTGCTGCCCGGGTAAAGCATCGATGCGGATGCCAA 3286  
 QY 241 agattcccgacattcttcacatccgcgcgttggagccacatcgcaagtggacaagc 300  
 Db 3287 AGATTCCCGACCTTGCTTCCATCGCGCGGTGGAGCCACCATGCGAGCGAGCGC 3346  
 QY 301 tgcggaaacacttgagcttccggcgaaacttccacagctggttgacctctttgcccgc 360  
 Db 3347 TCGGAAACCTTGAGCTTCCGCGGAACCTCCACGCTGTGTGATCACTTTGGCCGC 3406  
 QY 361 catccgcaagaacacccggaacttgcttcgcgcgtggaagaaggaattccggtatccg 420  
 Db 3407 CATTCGCCAGACACCCGGAACCTTGCTGCGACATGGAAGGACATTCGGATTATTCG 3466  
 QY 421 tgcgtccgacatcttggtggcgaattgctggaaggtccaccaggtcttgatcgcggtac 480  
 Db 3467 TCGCTCCGATCTTGTTGGGGAATTCCTGGAAGGCTCCACCCAGGCTTGTATGCGGATAC 3526  
 QY 481 ccacgtaagaactccacacactcatgtctgtgtagcatgacggcgaggtatgga 540





Kotonay, J.F., Nelson, W.C., Unayem, L.A., Ermolaeva, M.,  
 Salzb erg, S.L., Delcher, A., U tterback, T., Weidman, J., Khouri, H.,  
 Gill, J., Mikula, A. and Bisnal, W.  
 Whole genome comparison of Mycobacterium tuberculosis clinical and  
 laboratory strains  
 Unpublished  
 2 (bases 1 to 20400)  
 Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O.,  
 Peterson, J., Deloy, R., Dodson, R., Gwinn, M., Helt, D., Hickey, E.,  
 Kolonay, J.F., Nelson, W.C., Unayem, L.A., Ermolaeva, M.,  
 Salzb erg, S.L., Delcher, A., U tterback, T., Weidman, J., Khouri, H.,  
 Gill, J., Mikula, A. and Bisnal, W.  
 Direct Submission  
 Submitted (25-APR-2001) The Institute for Genomic Research, 9712  
 Medical Center Dr, Rockville, MD 20850, USA  
 Location/Qualifiers  
 1..20400  
 /organism="Mycobacterium tuberculosis CDC151"  
 /strain="CDC151"  
 /db\_xref="taxon:83331"  
 /note="clinical strain"  
 221..309  
 /gene="trna-Leu-4"  
 /product="trna-Leu"  
 221..309  
 /gene="trna-Leu-4"  
 complement(427..744)  
 /gene="MT200"  
 complement(427..744)  
 /gene="MT200"  
 /note="identified by Glimmer2; putative"  
 /codon\_start=1  
 /transl\_table=1  
 /product="hypothetical protein"  
 /protein\_id="AAK6448.1"  
 /db\_xref="GI:1381885"  
 /translation="MTRRLRVHNGVEDFEAFSYDAAPDIDRLYNLVDVATKR  
 IPQAPNFAFLPKFKRHRIYLPFRYYVAVRTTDEAIDILAVRGMENPNAVEAISGR  
 TFE"  
 complement(741..956)  
 /gene="MT201"  
 complement(741..956)  
 /gene="MT201"  
 /note="identified by Glimmer2; putative"  
 /codon\_start=1  
 /transl\_table=1  
 /product="hypothetical protein"  
 /protein\_id="AAK6448.1"  
 /db\_xref="GI:1381886"  
 /translation="MYVRRALSLVSLSRDEQILEYEHINGLAEAGHISEANQALI  
 EAKRADDHASTIDDEKIRARLG"  
 1211..2269  
 /gene="MT202"  
 1211..2269  
 /gene="MT202"  
 /gene="MT202"

gene  
CDS  
/trans\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AKK46486.1"  
/db\_xref="GI:1381887"  
/translation="MEAPYASGPTFERLRKSFQADLLPELOAGVHTIATVEAAD  
PAENESLLTAHHDMIAKIVGNYPLADEDETESSTGHNRDRDSWRDLRCGLP  
PGCHQPLVVGIVGQCEPMRPMPSGFLRPTPTRRSDRDLGRLADELASRGSD  
RLVLGIANGSGVPGVMEVSAALCAELDLVFLKGVQWMLGLMGLASGGVYND  
VSLSLRTDOOVRAIIBSETAELORELLAYRCGRVYDPPARVILVYDGIATASNL  
AAVATIRATGSESIYVAVPVGPTACRELAEMADVDVCAIMPAFEARGVYNDHFV  
TDDVSEVLLAPPTCAAT"  
complement(13399..2755)  
/gene="MT203"  
complement(13399..2755)  
/gene="MT203"  
/note="identified by Glimmer2; putative"





RESULT	7
MTCT270/c	
LOCUS	37586 bp DNA
DEFINITION	03-AUG-2001
	Mycobacterium tuberculosis H37Rv complete genome; segment 96/162
	BCI

ACCSSION	295388	AL123456
VERSION	295388.1	GI:3261759
KEYWORDS		
SOURCE		
ORGANISM	Mycobacterium tuberculosis H37Rv.	
REFERENCE	Mycobacterium tuberculosis H37Rv.	
AUTHORS	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium tuberculosis complex.	
	1 (bases 1 to 37586)	
	2 (bases 1 to 37586)	
TITLE	Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence	
JOURNAL	Nature, 393 (6685), 537-544 (1998)	
MEDLINE	98295987	
REFERENCE		
AUTHORS	2 (bases 1 to 37586)	
TITLE	Parkhill, J.	
JOURNAL	Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk	
COMMENT	On Jun 27, 1998 this sequence version replaced gi:2104310.	
	Notes:	
	Details of M. tuberculosis sequencing at the Sanger Centre are available on the World Wide web.	
	(URL: <a href="http://www.sanger.ac.uk/Projects/M.tuberculosis/">http://www.sanger.ac.uk/Projects/M.tuberculosis/</a> )	
	CDS have been renumbered from the original consuld submissions but the old gene designations are in brackets after the new gene numbers.	
	Gene prediction was based on a Hidden Markov Model of TB genes implemented in Tbparse (Kirogi) supplemented with visual inspection of positional base preference in codons, especially where there is an increase in the observed/expected third position G + C.	
	CAUTION: In some cases we may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtc, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-11bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.	
FEATURES	Location/Qualifiers	
source	1..37586	
	/organism="Mycobacterium tuberculosis H37Rv"	
	/strain="H37Rv"	
	/db_xref="taxon:83332"	
	/clone="Y270"	
gene	complement(51..854)	
	/gene="cysQ"	
CDS	complement(51..854)	
	/gene="cysQ"	
	/gene="cysQ"	
	/note="R2131C. (MTCY270.37). len: 267. cysQ, equivalent to M. leprae ctsQ.MYCLE P46726 cysQ protein homolog. (289 aa). Contains inositol monophosphatase family signature 1 (PF006029), significance uncertain. FASTA best: cysQ.MYCLE P46726 cysQ protein homolog. (289 aa) opt: 1374 z-score: 1597.0E(1): 0 (77.3% identity in 264 aa overlap)"	
	/codon_start=1	
	/transl_table=11	
	/product="cysQ"	
	/protein_id="CAB08639.1"	
	/db_xref="GI:2104347"	
	/db_xref="SWISS-PROT:O06244"	
	/translation="VSPSAKPDITFDITLDEADLADLADGKLLIVRAEIGFDQPMWT LGSAGRANSLLRQLRQAPPDGDAVLAEEADDLARKSDRWIITIDLGRTSESEFVGRDPAVLAIRMSNGEITDAVLAIPAGNVYRTQVPSGAAPGVGKRTLRATIA VSRTPAVLHIRQTLAIQPVISIGSNGAKKAAVVDGVYATVHAGGQEMSPAGAPK	

```

misc_feature
      VMLAGMHASRLDGSPLRYNODLPFLPDLMLCRFAVAPILLGAIADAMF"
      complement(552..593)
      /gene="cysQ"
      /note="P500629 Inositol monophosphatase family signature
      1"
gene
      945..1175
      /gene="Rv2132"
      /note="Rv2132"
      945..1175
      /gene="Rv2132"
      /note="Rv2132, (MTCY270.36c), len: 76. Function unknown
      but some similarity to Mycobacterium tuberculosis
      AL021924|MT020_4 (84 aa). FASTA score: opt: 142 z-score:
      210.0 E(): 0.00036; 41.8% identity in 55 aa overlap"
      /codon_start=1
      /transl_table=11
      /product="hypothetical protein Rv2132"
      /protein_id="CAB08638.1"
      /db_xref="GI:2104346"
      /db_xref="SPTREMBL:006243"
      /translation="MTTYSLADVAANAORLRKRENSIGLSEAVNELIRAGLTRQVA
      NRRQQTIDMGESIDYSNIGDALETLDGPASG"
      complement(1385..2173)
      /gene="Rv2133c"
      complement(1385..2173)
      /gene="Rv2133c"
      /note="Rv2133c, (MTCY270.35), len: 262. Function: unknown
      but equivalent to hypothetical M. leprae protein, Q49774.
      FASTA best: Q49774 B2126.C1.150. (262 aa) opt: 1447
      z-score:1647.5 E(): 0; (79.0% identity in 262 aa overlap)"
      /codon_start=1
      /transl_table=11
      /product="hypothetical protein Rv2133c"
      /protein_id="CAB08660.1"
      /db_xref="GI:2104345"
      /db_xref="SPTREMBL:006242"
      /translation="MLADGELTVLRIRASNAATFICESTIGLRSLHCYVYVGEER
      LMPDPGTLAGEISAVLYSTOLGMYLPFTITTRDPASIGMLQWQDPDADYDIP
      LGGPDVLDLPAPHRPRGYLYVLAHYDAGDEVYLNHADDITLRMAVFDVLINNAAR
      KGHITICGIDGOVYVNGDLCLAHYENKRLRVLMGMAKGRKIDQITLQAAVAGLADALGSP
      LAERLAGRIAAAEICGLAKRAOSLLDQPVMPGNGRPIPMPAF"
      complement(2184..2771)
      /gene="Rv2134c"
      complement(2184..2771)
      /gene="Rv2134c"
      /note="Rv2134c, (MTCY270.34), len: 195. Function: unknown
      but equivalent to hypothetical M. leprae protein, Q49789.
      FASTA best: Q49789 B2126.C1.228. opt: 1192 z-score: 1493.1
      E(): 0 (91.1% identity in 192 aa overlap)"
      /codon_start=1
      /transl_table=11
      /product="hypothetical protein Rv2134c"
      /protein_id="CAB08659.1"
      /db_xref="GI:2104344"
      /db_xref="SPTREMBL:006241"
      /translation="MARAIHVFRTPDHFVAGTGVGQNGTFVYLAHVHDSRVVYLEK
      OQVAVALERIGALFLFVNRRGCPVPEPEIDSLPLMFDAFERGIGLDMWSE
      AOSVVELAIVDAEFDASTVLDTEGPPAVAVFLLPSASOKAFATRSYVLSAGRP
      CPCLDEPLDPESHICARTNGTRDVLGSSDDAG"
      complement(2776..2780)
      /note="possible RBS, GGAGG, for Rv2134c"
      complement(2835..3545)
      /gene="Rv2135c"
      complement(2835..3545)
      /gene="Rv2135c"
      /note="Rv2135c, (MTCY270.33), len: 236. Function: unknown
      but equivalent to hypothetical M. leprae protein, Q49773.
      FASTA best: Q49773 B2126.C1.148 opt: 1183 z-score: 1390.9
      E(): 0; (74.8% identity in 236 aa overlap), also similar
      in C-terminus to Pmg2_Ecoli P36942 probable
      phosphoglycerate mutase 2 (215 aa) fasta scores: opt: 212
      z-score: 257.4 E():2.5e-07 27.9% identity in 190 aa
      overlap"
      /codon_start=1

```

```

      /transl_table=11
      /product="hypothetical protein Rv2135c"
      /protein_id="CAB08658.1"
      /db_xref="GI:2104343"
      /db_xref="SPTREMBL:006240"
      /translation="MTVILNHRNSTNTAGVLAGSGVDLDEKREGATGLIDRID
      LPIRAASPMILRCORVEPLAEALCLEPLIDREFSEVYGEWTRKIGDLYDEPLRW
      VQAPHSAAVFPFGGEGGLAQVOTRAVAAVRHDLADQGHVWLACTHDVIAKAVI
      ADAPGHMIDSFORTADPGSVSVRYTQLPFVLVHNHGARLAPALQAAASAGASP
      EPAVAVPDGAIVIGGSTD"
      complement(3542..4372)
      /gene="Rv2136c"
      complement(3542..4372)
      /gene="Rv2136c"
      /note="Rv2136c, (MTCY270.32), len: 276. Function: unknown
      possible membrane protein very similar to hypothetical M.
      leprae protein Q49783. FASTA best: Q49783 B2126.C2.190.
      opt:1023 z-score: 1259.2 E(): 0; (82.4% identity in 187 aa
      overlap) similar to BACA_ECOLI P31054 bacitracin
      resistance protein (273 aa) opt: 477 z-score: 590.1 E():
      7e-26, (35.6% identity in 267 aa overlap)"
      /codon_start=1
      /transl_table=11
      /product="hypothetical protein Rv2136c"
      /protein_id="CAB08657.1"
      /db_xref="GI:2104342"
      /db_xref="SPTREMBL:006239"
      /translation="MSMMQVIVLAAAGLTFEPLVSSGHLAIYRIFPSGDGASFT
      AVSOLGTEAAVYIYFARDIYIILSAMHLGVYKARNDRYLKAWYIIGTPTICIGL
      FPKDIDSGENVLMVYVYALVFSVYALAEVYKARNDYRLKAWYIIGTPTICIGL
      IYGVSRSGSTISAGTICGLDRELARFGLAIIVAVRSGLSPSPDHFHYEGNSA
      TGPOLLVATLIAFLVGLTRAVAAVLRFLVRNMTWVGIVLGVMTLVLATGVAT"
      complement(4436..4849)
      /gene="Rv2137c"
      complement(4436..4849)
      /gene="Rv2137c"
      /note="Rv2137c, (MTCY270.31), len: 137. Function: unknown
      C-terminus is very similar to hypothetical M. leprae
      protein B2126.C2.188 (150 aa). FASTA best: Q49782
      B2126.C2.188 (150 aa) opt: 469 z-score: 623.7 E():
      9.6e-28; (77.2% identity in 101 aa overlap)"
      /codon_start=1
      /transl_table=11
      /product="hypothetical protein Rv2137c"
      /protein_id="CAB08656.1"
      /db_xref="GI:2104341"
      /db_xref="SPTREMBL:006238"
      /translation="MRNKKSTSHSESGKLISSCGRPREMVLQVSLGMYTADNLT
      ADKREFAVEDISTGIFASGVGVGGRGSEFPHIRSLVYEITRPVADGPVQADYV
      VMAVRGLVIDLDLDESLAAAVRDSVAALAVSR"
      4864..5940
gene
      Query Match 24.8%; Score 392.2; DB 1; Length 37586;
      Best Local Similarity 56.4%; Pred. No. 2.1e-84;
      Matches 792; Conservative 0; Mismatches 608; Indels 5; Gaps 3;
      137 gatctgcgcgcgtcacctcgcgtatgctggcggaacgcggaatgctggtgcgga 196
      19632 GATCTGCGCGGGGTGCACATGCTCGCATCGCGGAGCTGCATGTCGGACATCCCGCA 19573
      197 atctctgctccgcggtgaagacagtcagtgttcgcgaagcaagaatcccgacattg 256
      19572 ATCTGTGTGACCGCGCGCGGTGCTCTCCGCGGCACACCCCAAGGAGTCCGCGGTGTG 19513
      257 ctccactccgcgcgctggagacacacatcgcaatggagacacgcctcggaataacttgag 316
      19512 CATGCGCTGCGGGCGGGCGCGGCGTGTATCCGAGTCCGACACCGCGTGTGAC 19453
      317 ctctccggaactccacacgctgctgtgtaacctcttttgcgcgaatttcgacaagaac 376
      19452 CTGTTGCCCGGCGGCGCGCGGTCTGCTACTACCATGCGCGCATCCCAAAACCAAC 19393

```

OY	377	ccggaactgtgtctgtgaacgttgaagaaggaattccggattatctgcgtctcagactgtt	436
Db	19392	cccgagctctgtcgaacggagagggccggcgacatttccctggctcggccggccgtctc	19333
OY	437	ggcgaattgtctgtgaaggtctccaccgaagttcttgatcgtggtatccaccgyltaagctcc	496
Db	19312	ggccaaattgattgacggcgccgacacacattgattggtctaccggacacccagacaaacac	19273
OY	497	accacactatgtctgttgaactatgtcagaacagcgtgacatgtgacaccagaattgtctac	556
Db	19272	acgagatgtccatgtcgtatcgttcccttcgacgacactgcgggcttgaccctcttgccgctc	19213
OY	557	ggcgcagacgtcccaacaaagcctctgacacaaatgcgcacatggaactggtgtgaatttacc	616
Db	19212	ggcggatgacgtcgggggagggccggtacacaccccatatccacgactgaggactgtttgtc	19153
OY	617	gtctgaagcagatgaattctgaacgcatcgtctgtctgcgtacaaagccaatgtgtcagttgctc	676
Db	19152	ggcgaagccgacgaamaacgattggctgctgtttgcagatgacacacccacgtgcgctgctac	19093
OY	677	accacatgtggaaccaagacacacactgtgactctctttaaacccttgaagcttaattccaaagt	736
Db	19092	accacacatcgaagttccgattaccctgacattcttgcagatgacacacccacgtgcgctgctac	19033
OY	737	ttcgcagattctcgaagagacgcatctaccccggaacggaactgtgtgttgcctgaagat	796
Db	19032	tttcgactccttctgtgagagccactgttcccgggggctcctgtgtgtgtgactgacgacac	18973
OY	797	cctcagcagcgtgagctgtggtggaagagtctgtctccgaaggtatcaagaactgttgttacc	856
Db	18972	cccgagggggccgctgcgtacgctgacggccg-cactgacatctggaaatttcagatgctgcctaa	18914
OY	857	ggtaacgctgtgacgcgtacagagacacacactgtgattctcagcgtatgtaccacatcgtgtgat	916
Db	18913	cgggtgcgtgctggcggtgtagacacattgacgacacactgtgtgtgtgtgtgacgaacgggggt	18854
OY	917	tcccaagttgtcgaagagaagaaacccgcgcacataaatgtatgtgacagaagaattctc	976
Db	18853	cggccgctccacacatttccggttgacctcagaaactgaccacacacacaggtcccgacgt	18794
OY	977	g-tgaattctcaaatccctgtgtgatacagtgtaactaaacggtgtgacgcgcctgtctgac	1035
Db	18793	gattggcgctgctgcgtgctggcgccgacacacattgacctacacggctggagacgctgctggcg	18734
OY	1036	cgaataacattgtgtgtgtggaagcgtctgacaagaactgtgtgaagcctgtgtcgtatctctcgg	1095
Db	18733	ccgggtgacgattgctggggcccgccggcagagactcttgacgggctgcggccttcgaaagc	18674
OY	1096	cgtgtgaaagcgcgcctttagtttccaaacgtgtctatcgaagtcggtcaaatlaatgtgcgtgc	1155
Db	18673	actggtggcgacgatttcgacactgctgttggaacctgcggcgctggaagggcgtgcgtgcgcg-	18613
OY	1156	tattatgatgatgacgcacacaccccaacggaagatgaactgtcagtgatcagctgtgcgcg	1215
Db	18614	--tgttccattactacacccacacacccacggaatgacgacacattgacggcgccggc	18557
OY	1216	caacccggtgtgaagccgcctctggaagaagggcgcgtgatactgcgtgtccacacaaattata	1275
Db	18556	cattggtgctcgaacagagggcagcggcgccctcgtatgctgtgtgtttcaaccctatttgta	18497
OY	1276	ctcagcaccatgaattccaaaaagagttctgcggtgggaactgtcactgtgcagacgctgc	1335
Db	18496	tttcggcgacaaaggacattgcgtcgtgactgttgggcgcgtgcgaattgccctgtgacgaggt	18437
OY	1336	cgtgtgtgtgtgaatttaccggagcgcgtgaaacaaacgcgtgtgatatgtctcctcgaat	1395
Db	18436	gtttcctgatactacactcagcagacgtcgtgaacaaacgcgtgcggctgtcacaggagcag	18377
OY	1396	catcaacgaatgcgtatgacattccacgtcgtgtgtgaataacttaattctctgtgaagccaga	1455
Db	18376	cctgcctgtagacactcgtgtgcgatacctgactgtgtccggaatttttggcgcgctgcacaa	18317
OY	1456	acgcatctgcagaataatgcagaagcactaatgatcatcgtgtctaccatgtgtgcagtttcgt	1515







Db 5825 ATCTGTAGACCGGGCGGTCTGCTGCTCAATGCCAAGAGTCGGTGTGTG 5884  
 QY 257 cttccactcgcgcgcgtggagaccacatcgacgtggagaaagctcggaaaccttag 316  
 Db 5885 CACGACATGGCGGGCGCGCCCTGATCCGA-TCGACATATATCGTATGCTGAC 5943  
 QY 317 cttccgcgaactcccaacgctggtgtagacctcttttcgcacatcccgcaaacac 376  
 Db 5944 TTGCTGCTGGCGCGCTACCCGACATCAATCCGACATCGCATCTCCCAAGACTAC 6003  
 QY 377 ccggaactctgctgcacgtgaaagaagcattccggtttatgctgcacgtctgtg 436  
 Db 6004 CCGAGCTGTGCAAGCCCGGAGGGCGGCAATTCGGTCTTATGCGATGGCCGTGTG 6063  
 QY 437 ggcgaattctgcagagctccacccaggtcttgcggtggtcccaagtaaacctcc 496  
 Db 6064 GCCAGGCTGATGGATGGGTACACACGCTGATGGTCTTCGACACACGGCAAGACTACG 6123  
 QY 497 accacccctatgtctgtgtagctatgtagcagcagcgagcattgacagcttctatc 556  
 Db 6124 ACGACGTGATGCTGCTGTTGCTGCTGACGACATGATGATGATGATGATGATGATG 6183  
 QY 557 ggcgcgaactcacaagcgtgcacacatgacacatggaactggtgaggtcttacc 616  
 Db 6184 GCGGTGAGCTTCCGCTGGTGGGACACACCCGACCCGCTCGCGCTGCTGCTGCTT 6243  
 QY 617 gctaaagcagatgaatcagcagcagcagcagcagcagcagcagcagcagcagcagc 676  
 Db 6244 GCGAGGCGCAGCAAGAGATGGCTGTTACTGGAATACAGCCCAACTACGCGTGTG 6303  
 QY 677 accaatgtgaaccac 736  
 Db 6304 ACCAATCATCATCCATCTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 6363  
 QY 737 ttccagattctcagagcagcagcagcagcagcagcagcagcagcagcagcagcagc 796  
 Db 6364 TTGACATCTTCTGAGCGATTCGACACTGGGGGAGACCTGCTGCTGCTGCTGCTGCT 6423  
 QY 797 cctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 856  
 Db 6424 CCGTGGCG 6483  
 QY 857 ggtacgcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 916  
 Db 6484 GGTCTCCGACACCGATAGTACGATCTTGGCCCGCGCGCGCGCGCGCGCGCGCGCG 6537  
 QY 917 tcccaagctgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 976  
 Db 6538 CAACAGGCGACCGGACCGGTCCGACATCCAGTGGCTGCTGCTGCTGCTGCTGCTGCTG 6597  
 QY 977 gttattcttaaaatccctgtgtatcaatgtagtactcaagctcagcagcagcagcagc 1036  
 Db 6598 ATGGGCTTACCGGCTGCTGCGGACATGCGCGCTCATATCGCTAGCTGCTGCTGCTG 6657  
 QY 1037 ggaatccctgt 1096  
 Db 6658 GCGATTCAGATGTGTGCTGCGACGATGAATGCTGACGCGGTGCTTCAAGGGGT 6717  
 QY 1097 gtggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1154  
 Db 6718 GTGGGTGCGCGCTTCAATGCTGTGGACTTCAAGAGTCCGCGCAAGCGCTCAAACTCG 6777  
 QY 1155 ct----attatgatgatcagcacacacacacacacacacacacacacacacacacac 1210  
 Db 6778 GTGGGGGTGTTGACGACGACTACACACATCCAAAGGATCAGCGCGGTACGCGG 6837  
 QY 1211 ggcgcac 1270  
 Db 6838 TTCCGTATATCTTACAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6897  
 QY 1271 ttatactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1330

Db 6898 CTATATTCGGGCAACAGGCTTTGCCAGAGATTGGTGTGTCCTGATGCTGCCGAT 6957  
 QY 1331 gctgcgcgt 1390  
 Db 6958 GAGGTGTATGTTGT 7017  
 QY 1391 gaaatcaccagatgcga-tgaccattccagctgtgtgtgtgtgtgtgtgtgtgtgtgt 1449  
 Db 7018 GCTAGT 7077  
 QY 1450 cccagacagcattgagaagaatccagagcagcagcagcagcagcagcagcagcagcagc 1509  
 Db 7078 GCTCATAGAGT 7137  
 QY 1510 ttcctgtaccatgt 1541  
 Db 7138 CGACGTGGCGCTGTGGGGCGGAGATCGTTG 7169  
 RESULT 11  
 AB003132  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 AB003132 4116 bp DNA BCT 04-AUG-1997  
 Corynebacterium glutamicum gene for MurC, FtsQ, FtsZ, complete cds.  
 AB003132  
 AB003132.1 GI:2308989  
 FtsZ, FtsQ, MurC.  
 Corynebacterium glutamicum DNA.  
 Bacteria; Firmicutes; Actinobacteria; Actinobacteriales; Actinomycetales; Corynebacteriales; Corynebacteriaceae;  
 Corynebacterium.  
 1 (bases 1 to 4116)  
 Kobayashi, M.  
 Direct Submission  
 Submitted (15-APR-1997) to the DDBJ/EMBL/GenBank databases. M1K1  
 Kobayashi, Mitsubishi Chemical Corp., Tsukuba Research Center; Ami  
 8-3-1, Inashiki, Inashiki, Inashiki, Inashiki, Inashiki 300-03, Japan  
 (E-mail:3709292@cc.m.kagaku.co.jp, Tel:0298-87-1011,  
 Fax:0298-87-3259)  
 2 (sites)  
 Kobayashi, M., Asai, Y., Hatakeyama, K., Kijima, N., Wachi, M., Nagai, K.  
 and Yukawa, H.  
 Cloning, sequencing, and characterization of the ftsZ gene from  
 coryneform bacteria  
 Biochem. Biophys. Res. Commun. 236 (2), 383-388 (1997)  
 97382442  
 FEATURES  
 source  
 1..4116  
 /organism="Corynebacterium glutamicum"  
 /db\_xref="taxon:1718"  
 1..273  
 /gene="murC"  
 <1..273  
 /gene="murC"  
 /codon\_start=1  
 /transl\_table=11  
 /product="MurC"  
 /protein\_id="BA21685.1"  
 /db\_xref="GI:2308990"  
 /translation="EF0KEFALSLADAAVLEIYGARQPVQVSSIIITDAMTIP  
 VYVEPNFSAVERIAIADPNDIVLTMGAGSVTLAEIILDOLNN"  
 301..969  
 /gene="ftsQ"  
 /gene="ftsQ"  
 301..969  
 /gene="ftsQ"  
 /codon\_start=1  
 /transl\_table=11  
 /product="FtsQ"  
 /protein\_id="BA21686.1"  
 /db\_xref="GI:2308991"  
 /translation="NNKQVIAIVGVVVVLAIVGVAVFPIIKVGNIEVGTATRD  
 PDQVLEVSGIVEENLFRVDAQAQNTVELPWSVTVNRAALPTITVELTEREPAV  
 FIKRADGVIDITGEKEIIIGTPVGTVEVSGADEGSEVLPVAVIVNAKAQD"

```

TESTVVEAPDPOFDILKMKNDGREIYWGSSNNHDKAVANSTVLKKGGRWNISPSM
VTVR*
1220..2536
/gene="fts2"
1220..2536
/codon_start=1
/transl_table=11
/product="fts2"
/protein_id="BAA21687.1"
/db_xref="GI:2308992"
/translation="MSTPNNYLAKIKVVGGGGDNANRMEIEGLKVEEFAVNTDS
QALMFSDADVKLIDIGREATRGAGANPEVRAAEYKHEIEETIKGADVFYAGE
GGAGTGAAPVYRIAKMGALIGVYTPPEFEGRRTROAEGTALKEVCDLIY
IPNDRLLEGDANLSIMEAFRAADINLVINGOGITNLITPICINVDPAVRSVMSAG
SALMGVSGARDNRVYATEQAINSPILEAEPADCTOVLISPGSGDGLMVYNAAS
MVRERSEVDNVLIFGIIIDNDGDEVRYANGFDPARSAENRRGITPAFPAAP
VQOOOYPTNATTPPKESITGGAKEEDPILSSAGARRIEETRSGGGLFTTGNDR
DYRDERREDRDERNDRSTDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDRDR
2564..3304
/note="hypothetical protein similar to YFIH"
/codon_start=1
/transl_table=11
/protein_id="BAA21688.1"
/db_xref="GI:2308993"
/translation="MDSIDPRNRPYRYETTRAGVSSQSPYATFNLCGHVYPRANA
SNRRRLADITGLSPKYYMEQIHSNTVYIDEAPGQAVETDLYTTRGALAV
LYADCPVILSTDDGVIAVHAGMGRNGIYAKTIKHEELGAKPSRIHALMGAA
SGAKTEPPEMAADVEAKIPSLAKITGTGIDIRAGLLRMLSLGQVOMIDSPRCT
IEEDLESTYRREGTGRQAGVWMLPKKA"
3345..4010
/note="hypothetical protein similar to YPII or YPI5"
/codon_start=1
/transl_table=11
/protein_id="BAA21689.1"
/db_xref="GI:2308994"
/translation="MARIDATLNEHNRPGSVRLPVKHEPVEDIKIIOELGIVAV
GENRQDARAKALELPDMDFHIGIQIOSKANSIARMSAAVSDSEKIAELGRVYA
LALEDRTSDLEPCFIQLSLDGPSSRGTPLSQVTLADCSLPTTHRFELGACVPP
LGMDPKAFSQAADVLSGIEEHFDRSLSEFSGMSGDLVAIAIKHGSTIVRGTEILGR
PLA"

BASE COUNT      1009 a      1106 c      1148 g      853 t
ORIGIN

Query Match      18.3%  Score 289.8;  DB 1;  Length 4116;
Best Local Similarity 99.3%;  Pred. No. 1.5e-59;
Matches 291;  Conservative 0;  Mismatches 2;  Indels 0;  Gaps 0;

QY 1289 gaattccaaagagatcgcggggacacgtacacgtgcagacgctgcgtggtcttgcg 1348
|||||
1 GAATTCCAAAGGAGTTCCGCGGAGGACGCTACGCGAGACGCTGCCGTGCTCGAG 60

Db 1349 atttagcgagcgccggaacacgcgtgagcggtgctctcgaaatcattaccgattgcg 1408
|||||
61 ATTTCGGAGCGCGCAACCGGTGATGCGCTGCTCGGAAATCATCACCGATGCG 120

QY 1409 atgaccattcagtggtacgaacctaatctctcgtcagctcccaagaacgattgcaga 1468
|||||
121 ATGACCATTCAGTGTGTGACGAACCTAATTTCTCTGCAAGTCCAGAACGCAATGAGAA 180

Db 1469 atcgcgagacctatgacatcgctcaccacgaggttggaagttccgtagacatgcttgc 1528
|||||
181 ATCGGAGAGCACTAATGACATCGTCTCACACGAGGTGAGGTTCCGACACGATGCTGCT 240

QY 1529 ccagaatactgatactgacgtcgaacaaataggaagcgtgaagtgaaacaagcgacg 1581
|||||
241 CCAGAAATCTGTGATCAGCTGCAAAACATTTAGGACGTAAGTGAACAAGGACG 293

Db 241 CCAGAAATCTGTGATCAGCTGCAAAACATTTAGGACGTAAGTGAACAAGGACG 293

RESULT 12
BLFITSZ      5546 bp      DNA      BCT      19-JAN-2000
LOCUS

```

```

DEFINITION      Brevibacterium lactofermentum murC(partial), ftsQ, ftsZ genes
ACCESSION      ORE5, ORF6 DNA.
VERSION      Y08964
KEYWORDS      Y08964.2 GI:6723445
SOURCE      cell division protein; divB gene; division initiation protein; ftsQ
            gene; ftsZ gene; murC gene; putative vak 1 protein; putative YPT5
            protein; UDP-N-acetylmuramate-alanine ligase.
ORGANISM      Corynebacterium glutamicum
            Corynebacterium glutamicum
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Corynebacteriales; Corynebacteriaceae; Corynebacteriaceae;
            Corynebacterium.
REFERENCE      1 (bases 1 to 5546)
AUTHORS      Honrubia, M.P., Fernandez, F.J. and Gil, J.A.
TITLE      Identification, characterization, and chromosomal organization of
            the ftsZ gene from Brevibacterium lactofermentum
JOURNAL      Mol. Gen. Genet. 259 (1), 97-104 (1998)
MEDLINE      98409265
REFERENCE      2 (bases 1 to 5546)
AUTHORS      Gil, J.A.
TITLE      Direct Submission
JOURNAL      Submitted (21-OCT-1996) J.A. Gil, Universidad de Leon,
            Microbiologia, Campus de Vegazana S/N, 24071 Leon, SPAIN
REMARK      Revised by [3]
REFERENCE      3 (bases 1 to 5546)
AUTHORS      Gil, J.A.
TITLE      Direct Submission
JOURNAL      Submitted (24-MAY-1999) J.A. Gil, Universidad de Leon,
            Microbiologia, Campus de Vegazana S/N, 24071 Leon, SPAIN
COMMENT      On Jan 20, 2000 this sequence version replaced gi:1769558.
FEATURES
            source
            gene
            CDS
            /gene="murC"
            /locus_tag="murC"
            /NC_number="6.3.2.8"
            /codon_start=1
            /transl_table=11
            /product="UDP-N-acetylmuramate--alanine ligase"
            /protein_id="CAA70160.1"
            /db_xref="GI:1769559"
            /db_xref="SPTREMBL:P94335"
            /translation="KEFAEALSLADAAYVLEIYGAREQPDVGSSSEIITDAMTIPVY
            ENFSAVPERIAELIAGNDIVLMGAGSVTMLAPILDOLNN"
            292..960
            /gene="ftsQ"
            292..960
            /gene="ftsQ"
            /function="division initiation protein or cell division
            protein"
            /note="alternative gene name divB"
            /codon_start=1
            /transl_table=11
            /protein_id="CAA70161.1"
            /db_xref="GI:1769560"
            /db_xref="SPTREMBL:P94336"
            /translation="MNRKVAIVAGVVVLAALGVAMFVPLIKVGNIEVGTARTD
            PQQVLEVGIVGEGNLEFVDAAGONIVLEPWSVTVNRLSPSTIVLETEREPAV
            FIKRADGSHVIDEGKEIILIGTPVGTVEVSADGENSEVPAVIAVINAIKQDDAM
            TESIVVEAPDPOFDILKMKNDGREIYWGSSNNHDKAVANSTVLKKGGRWNISPSM
            VTVR"
            1211..2527
            /gene="ftsZ"
            1211..2527
            /gene="ftsZ"
            /codon_start=1
            /transl_table=11
            /product="cell division protein"
            /protein_id="CAA70158.1"
            /db_xref="GI:1769561"

```





```

gene      complement(1..447)
          /gene="SC9B1.01c"
misc_feature 1..103
          /note="Nominal overlap with Streptomyces coelicolor cosmid
          /note="SC9B1.01c"
CDS       complement(1..447)
          /gene="SC9B1.01c"
          /note="SC9B1.01c, partial CDS, possible integral membrane
          protein, len: 149aa; similar to many eg. TR:CA38499
          (EMBL:AL035636) putative integral membrane protein from
          Streptomyces coelicolor (830 aa) fasta scores; opt: 365,
          z-score: 428.6, E(): 1.5e-16, (44.7% identity in 141 aa
          overlap). Contains several possible membrane spanning
          hydrophobic regions."
          /codon_start=1
          /transl_table=11
          /label="SC9B1.01c"
          /product="putative integral membrane protein"
          /protein_id="CAB41547.1"
          /db_xref="GI:4691380"
          /translation="MSPIDARRRRPPRRDDLVASLVEFLVALPLICGVAVASGPAP
          LGLVTGVGLVGLVGLFPGSSLVQVSPAGLVLFVGVADQGLPALGVIVATGIVOL
          AMGALKRGWFRALISLVVEGLAGIGLVILAGQLYSAGLQAPASG"
          complement(456..1037)
          /gene="SC9B1.02c"
          /note="SC9B1.02c"
          /note="SC9B1.02c, probable carbonic anhydrase, len: 193aa;
          similar to many eg. SW:CYNT_ECOLI carbonic anhydrase from
          Escherichia coli (219 aa) fasta scores; opt: 512, z-score:
          618.8, E(): 3.8e-27, (43.6% identity in 188 aa overlap).
          Contains two Pfam matches to entry PF00484 Pro_CA,
          Prokaryotic-type carbonic anhydrases and Prosite matches
          to PS00705 Prokaryotic-type carbonic anhydrases signature
          2 and PS00704 Prokaryotic-type carbonic anhydrases
          signature 1."
          /codon_start=1
          /transl_table=11
          /label="SC9B1.02c"
          /product="probable carbonic anhydrase"
          /protein_id="CAB41548.1"
          /db_xref="GI:4691381"
          /translation="MQLPDNARMFGRRPEPTGHAEGSPVFLFICSDSRVVPALI
          TGAAPQLELRAGNIYPIPGTGHPSGEATVEYAVOVIGVADIVGCHSHCGACVAV
          LVAGDLTAVPVRMDLNAHADERKPCDDPADPVACAVOHVYLAQDLRLRSYPCVENR
          LADGRLRGWYEVHTCTVPAHNAHDADETL"
          complement(504..608)
          /gene="SC9B1.02c"
          /note="SC9B1.02c"
          /note="Pfam match to entry PF00484 Pro_CA,
          Prokaryotic-type carbonic anhydrases, score 31.50, E-value
          4.7e-08."
          /note="Pfam match to entry PF00484 Pro_CA,
          Prokaryotic-type carbonic anhydrases, score 31.50, E-value
          4.7e-08."
          complement(666..1019)
          /gene="SC9B1.02c"
          /note="SC9B1.02c"
          /note="Pfam match to entry PF00484 Pro_CA,
          Prokaryotic-type carbonic anhydrases, score 145.30,
          E-value 4.5e-41."
          complement(744..806)
          /gene="SC9B1.02c"
          /note="SC9B1.02c"
          /note="PS00705 Prokaryotic-type carbonic anhydrases
          signature 2."
          complement(915..938)
          /gene="SC9B1.02c"
          /note="PS00704 Prokaryotic-type carbonic anhydrases
          signature 1."
          complement(1116..1517)
          /gene="SC9B1.03c"
          /note="SC9B1.03c"
          /note="SC9B1.03c, possible secreted protein, len: 133aa;
          contains possible N-terminal signal sequence."
          /codon_start=1
          /transl_table=11
          /label="SC9B1.03c"

```

```

          /product="putative secreted protein"
          /protein_id="CAB41549.1"
          /db_xref="GI:4691382"
          /translation="MYRRTARTRVAVLAVFLALQFPAGSGFSATYNGDAVAHAR
          PGRIPSGTAPHDETCTCEGRCGCEPTAGCAGVDRRTTAASQSTPRPHPARASAPQ
          ARGSVASGRATGHPRRADAPAVLQVFC"
          complement(1707..2759)
          /gene="SC9B1.04c"
          /note="SC9B1.04c"
          /note="SC9B1.04c, possible ATP/GTP-binding integral
          membrane protein, len: 350aa; similar to hypotheticals eg.
          TR:P71967 (EMBL:Z80225) from Mycobacterium tuberculosis
          (369 aa) fasta scores; opt: 1129, z-score: 1182.3, E(): 0,
          (57.8% identity in 325 aa overlap). Contains possible
          membrane spanning hydrophobic regions. Also contains
          Prosite match to PS00017 ATP/GTP-binding site motif A
          (P-loop)."
          /codon_start=1
          /transl_table=11
          /label="SC9B1.04c"
          /product="putative ATP/GTP-binding integral membrane
          protein"
          /protein_id="CAB41550.1"
          /db_xref="GI:4691383"
          /translation="MSLCARPRVPADRLVAEMPPRPDSVREGTVPPNPQPSQSE
          AVRVLDEPAGIGKAPGTGGRRGFGRLGRAPTPGAPRVYLIDGYGVGKTHLLA
          SIMHATPAEPKAKAGTEFVELTNVGLAFQOQVOTLGSRLICIDEPELDDPGDVLV
          VESLARLVGVGVALAATSNLTGKGEFGFAADPLRIGLISAHFPAURLRIGEDYR
          HGLPEAPAPYSDQVQRAARATGEGSLIDFPALDHLARVHPSRYCALTDGLTAVLC
          TGVHPVPOSTRALVYLAURLVDREVPVLAGLFPDLRFEBMLKGRKKFRFRAIS
          RLTAARARARLVPTA"
          complement(2466..2489)
          /gene="SC9B1.04c"
          /note="PS00017 ATP/GTP-binding site motif A (P-loop)."
          /note="PS00017 ATP/GTP-binding site motif A (P-loop)."
          /note="High GC content."
          /note="High GC content."
          2863..3232
          /gene="SC9B1.05"
          /note="SC9B1.05"
          /note="SC9B1.05, possible hydrolase, len: 265aa; similar
          to hypotheticals eg. TR:P71968 (EMBL:Z80225) from
          Mycobacterium tuberculosis (258 aa) fasta scores; opt:
          536, z-score: 585.9, E(): 2.6e-25, (44.3% identity in 235
          aa overlap). Also similar to several proposed ribofavin
          deaminases eg. SW:RIBG_BACAM ribofavin-specific deaminase
          from Bacillus subtilis (371 aa) fasta scores; opt: 157,
          z-score: 177.0, E(): 0.015, (28.3% identity in 180 aa
          overlap)."
          /codon_start=1
          /transl_table=11
          /label="SC9B1.05"
          /product="putative hydrolase"
          /protein_id="CAB41551.1"
          /db_xref="GI:4691384"
          /translation="MRRLFPVSPATDPTAGPGCGEPTREWSLAELAAVPEPDAR
          RTWLRANWSTLIDGANOHDGRSQPISSAADRVRVGTIRALADVIVAGATVQDEYR

```

```

Query Match      14.7% Score 231.8; DB 1; Length 24800;
Best Local Similarity 51.8%; Pred. No. 1.3e-45;
Matches 747; Conservative 0; Mismatches 622; Indels 72; Gaps 7;

```

```

0Y 141 tgcctccgcgtccactcctcgtatctgctgagcagcaatgctgctgctgcaatcc 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4926 TGGACCGACCGCCACTCATCGGCATCGCGCGCGGATGTCGGCATCGCCAGATCC 4985
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
0Y 201 tgcctccgcgttaagacagctactgcttccgctgctgcaaatgcttccgcactctgctc 260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4986 TCGCCACGCGCGCGCGCCAGGTCGCGGACGCGCAAGAGTCTCGGACCGCGGACG 5045
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
0Y 261 cactccgcgcgttgagagaccacacgcagctgagagacagcttgagagaaaccttgacctt 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

D	5046	CGCTGGGCGCCCTTG6GGCCGACCGATGCAATG6GGAGCGGGGAGCACTTG-----	5093
Q	321	CGGCGAaattcccaacgctgtgtgtgaactctcttttccgccaatccgcaagaaccgg	380
D	5100	CGCGCAACCGAGAGTGGCTGTGTGTCTCT-----CGAGCATCCGAGGAATACTCG	5153
Q	381	aattgtctgtcaagctgtgaagaagaattccggtattatctgtctccgaactgttggcg	440
D	5154	AGCTGTTCGGCCCGCCCGACAGTGGGCAATCCCGTGTGTCAACGCTCCGAGCGCTTGCGG	5213
Q	441	aattgttgaaggtcccaacccaagttctatctgtcggtgtaccacaggtgaagactccaca	500
D	5214	CCTGTATGAACGGCTGTGGCCCATCGCGTGGCCGGCAACCAAGCAACCAACCA	5273
Q	501	cctctatgtctgtgttagctatgtcaggcaagcgagatgatcatcaagctttgtatgcg	560
D	5274	CTGTGATGTGGCGGTCTCTCCCTGTGCGACTGTGAACCTGGGCGCTGTCAATCGATCGGG	5333
Q	561	gaagaagtcacaagaagctgtgcaacaatgtgcacacatgtgaactgtgtgtatctatcgt	620
D	5334	GCACACTGTGACCCCCCGGCTCCAAAGCCCTGTCAAGCGGAGGAGGAGATCTTGTCGGC	5393
Q	621	aagcaataaatactcagcactgcgtctgtcgcttaagaagcaaatgttgcaggtgtacaa	680
D	5394	AGCGGAGCAAAAGCAGACCGGAGCTCCACAAATAGCTCCGAGAGTGGCATGTCTCTCA	5453
Q	681	atgtgtgaacacagacacactgtgaactctttaaacccttgaaagctacttccaagtctcg	740
D	5454	ACGTGCAAGTGTGACACCAACCACTCAACTCACTGTCAATGAAGAAATTAAGATCTCTTGC	5513
Q	741	acgaatttcgcaagaagcatcacaccgcaagcgcaagctgtgtgtgtctgtaaagatcttc	800
D	5514	AGACCTTTCGCGCGCAAGATGTGTCCGCGGGGCAACCTGTGT-----GATCGGG	5560
Q	801	acacagacggagagctgtgggaaggtctgtctcgcaagaaggtataaagactgtgtgtacgta	860
D	5561	GCGGACACAGAGAGGCGCCCGGAGACTGTACCCAGACGCTGTGAGAGCGCGGATCGAGAGG	5620
Q	861	cgcgtgtacgcgttacaggaacacacccctgaagtttcacgcaatgtaccatcgttgaattcc	920
D	5621	ACCTAGCGAGATCCGAGGAGCGCGACGTCACATCTGTCCGTGCTCCGCGACGGGCGCTC	5680
Q	921	aagttgtcgcaagaagcagcccgcgcaacatcaatgtatgtacaggaagatctctgtga	980
D	5681	AAG-----AGGAGAGTCAACGTGTGTCTGAGAGGCGCGAGAGCTGACCTTG	5726
Q	981	tcttccaatccctcggtgtgatacatatgtatctaagctgtgcagcgccctgtcgtgcgcat	1040
D	5727	CGGTCTCGGTCTCCGCGCGGCAATACGCCCAATAGCGCGCTGTGCGCGCGGGCG	5786
Q	1041	acctgtgtgtgtgtgtgagagctgcgaagaactgttgaagagcttgcgaatttctccgcgtgc	1100
D	5787	CGCCCGCTGGGGTCTCCCGCGCGCACTGTGCCCCCGGCGCTGACGCCCTACAGGGGCTCA	5846
Q	1101	gaagcgcgcttgaagttccacagctgtcatcagagggtggaattaaatgaagagctgtattt	1160
D	5847	AGCGGGCGCTCAAGCTCAAGGAGGAGGGCGCGCGGCTGCAGAGTCTGTGACTC-----	5898
Q	1161	atgatgatattacgacacaccacgaagaagtactgcaggtgtctcaagcgtgtgcgcacc	1220
D	5899	-----CTAGCGGCAACACCCACGAATATACCGCGACGTGAGAGGCAATGCGCGC-	5950
Q	1221	gggtgaagagcgctgtgaagagcgctgtcatcgttcgttcgcttccacaacacatatattactac	1280
D	5951	-----GCGGTGGGGAGACCCCGCATCTGTGTCTCTTTCAGCGGCAACTCTTCTCC	6002
Q	1281	gcaacataaattccaaaagattctcggtgggagccttcaactgtgcagagcgtctgtcg	1340
D	6003	GCACCCAGAGAGCTGGAGAGAAATGTGGGCAAGCCCTGTGGCCTTGAGACCTCGGTGTCG	6062
Q	1341	tgttgttagatttaacgagcgcggaacacacggtgtgatatgtcgtctcgtgaatacatca	1400
D	6063	TCTGTGACTCTACCCCGCCCGGAGAGACCCGATCCGGGGGTGTACCAAGAGCTGATGCG	6122

QY 1401 ccgattgcgattgacaccttcacgttggtgtacgaacctaattc-----tcgcagttcc 1451  
Db 6123 TCAGAGCGGGCCGCCGCCGCCGCCGAGGTACACCGCGTCCAGCAGAGAGCGCTTCCG 6182

QY 1452 caaagaacgattgcagaatactgcagagactaatgacatcgtgtccacatggtgcaggtt 1511  
Db 6183 CCGCGCTGGTGGCGGGAGATGGCGAAGCGCGGATCTGTTCTCACCATTGGCGCGGGG 6242

QY 1512 ccgttgaccatgcttgcctccagaatacccttgatcagctgcgaacaacttaagacgtaagt 1571  
Db 6243 ATGTCAACCGACCTGGAGCCCGCATCTCTGAGAGACTGTCCAAATGAGGGCGCTGAGACTC 6302

QY 1572 a 1572  
Db 6303 A 6303

RESULT	14
MLC266/c	
LOCUS	
DEFINITION	MLC266 36859 bp DNA
ACCESSION	AL022602
VERSION	AL022602.1 GI:3080471
KEYWORDS	3-deoxy-D-arabino-heptulosonate 7-phosphate synthase; anthranilate phosphoribosyltransferase; aroG; ctaG; cytochrome c oxidase subunitIII; ftsH; long chain fatty acid-coA ligase; murC; murD; murE; murF; murG; murX; pnpB; penicillin binding protein 2; phospho-n-acetylmuramoyl-pentapeptide-transferase; pseudogenes; serine/threonine protein kinase; trpD; udp-n-acetylmutamate:alanine ligase; udp-n-acetylmuramylalanine-d-glutamate ligase; udp-n-acetylmuramylalanine-d-glutamate-2,6-diaminopimelate ligase; Mycobacterium leprae
SOURCE	
ORGANISM	Mycobacterium leprae

REFERENCE AUTHORS TITLE	JOURNAL MEDLINE REFERENCE AUTHORS TITLE	COMMENT
1 (bases 1 to 38859) Elismeyer, K., Honore, N., Use of an ordered cosmid library to deduce the genomic organization of <i>Mycobacterium leprae</i> Mol. Microbiol. 7 (2), 197-206 (1993)	2 (bases 1 to 38859) Brown, D. and Churcher, C.M. Unpublished 3 (bases 1 to 38859) Parkhill, J., Barrrell, B. G. and Rasthream, M. A. Direct Submission Submitted (22-Apr-1997) Mycobacterium leprae sequencing project, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: barrrell@sanger.ac.uk Cosmids supplied by Dr. Stewart T. Cole, [3] unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France Requests for cosmids should be sent to Karin Elismeyer (ke@pasteur.fr)	Woods, S.A., Caudron, B. and Cole, S.T.

The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous. The length in codons is given for each CDS. Usually the highest scoring match found by fasta -o is given for each CDS. Details of the M. leprae sequencing at the Sanger Centre are available on the World Wide Web. (<http://www.sanger.ac.uk/Projects/LM3>) CDS are numbered using the following system eg MLCB33.01c. ML (M. leprae), CB33 (cosmid name), .01 (first CDS), c (complementary strand).

CDS which show significant similarity to other CDS in the database. The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon. All CDS over 100 codons have been analysed. Gene prediction is based on positional base preference in codons especially where there is an increase in the observed/expected third position G + C. CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon. IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid B268 is overlapped at the 5' end by MUB27CS and at the 3' end by MUB154CS/MUB155CS.

# FEATURES

```

source
1..38859
    /organism="Mycobacterium leprae"
    /db_xref="taxon:11769"
    /clone="Cosmid B268"
    /complement(1..705)
    /gene="murC"
    /note="murC"
    /complement(41..705)
    /gene="murC"
    /note="MUB268.04c, murC, probable
    udp-n-acetylmuramate-amine ligase, partial CDS, len:
    >231 aa; highly similar to many e.g. MURC-ECOLI P17952
    udp-n-acetylmuramate-amine ligase (EC 6.3.2.8) (491 aa),
    fasta scores: opt: 542 z-score: 704.5 E(): 4.1e-32, 41.98
    identity in 227 aa overlap"
    /codon_start=1
    /transl_table=11
    /product="udp-n-acetylmuramate-amine ligase"
    /protein_id="CA18667.1"
    /db_xref="gi:3080472"
    /db_xref="sptrembl:069551"
    /translation="MAGQLPELQVHMYGIGSGMSIARILLDRGLGSGDAKE
    SRVYHLRANGALIRIHGASLDLPLGATVATITRTAIPKTNPELVEARRGIPVL
    LRSAVALRLMDGCTTLMVAGHGTSTTSMVVAQHCCDPSFVGGELVAVTNAH
    HGIGACFVAEDASDLSLEYTPMVAVTVIDSDHIDPFYSVDAYIRVDFSEVERFAL
    GGLVVCNDPFGAALA"
    /complement(1..6446)
    /note="overlap with MUB27CS from: 1 to: 6444. There are
    sixteen conflicts between this sequence and ours. In each
    case our sequence has been checked, and is thought to be
    correct"
    538
    /note="conflict: this base is missing in MUB27CS"
    /complement(702..1925)
    /gene="murG"
    /complement(702..1925)
    /gene="murG"
    /note="MUB268.02c, murG, probable
    udp-n-acetylglucosamine-n-acetylmuramyl-(pentapeptide)
    pyrophosphoryl-undecaprenol-n-acetylglucosamine
    transferase, len: 407 aa; highly similar to many e.g.
    MURG-ECOLI P17443 (354 aa), fasta scores: opt: 461
    z-score: 727.7 E(): 0, 38.1% identity in 365 aa overlap"
    /codon_start=1
    /transl_table=11
    /product="udp-n-acetylglucosamine-n-acetylmuramyl-
    -(pentapeptide) pyrophosphoryl-undecaprenol
    n-acetylglucosamine transferase"
    /protein_id="CA18668.1"
    /db_xref="gi:3080473"
    /db_xref="sptrembl:069552"
    /translation="MNSVREPTKGRGSPVADALSYNPLSYVLGGGTAGHYEP
    AAVADALALDPOVRIALGTSKGLRVEBGYHLELITPVLPKXLTGLDARLP
    LRVNRAVRETRAVEVEAHVYVGGYVLAAGISIPVRRRIPIYVVAHMANRA
    GIANRGVTRAEVLSAVPGSGVGAEGVGLPHTITTLNRPALRDARKHGCFTDA
    ARVILVFGSGGAVSLNRNAYAGAEDLASGVAVLHAYGKNTLELTRPEYGEPTVA

```

```

RBS
    /note="possible RBS upstream of murC"
    1320
    /note="conflict: G is C in MUB27CS"
    1819
    /note="conflict: C is T in MUB27CS"
    1824
    /note="conflict: this base is missing in MUB27CS"
    /complement(1922..3526)
    /gene="ftsW"
    /complement(1922..3526)
    /gene="ftsW"
    /note="MUB268.03c, ftsW, probable cell division protein,
    len: 534 aa; similar to many e.g. FTSW-ECOLI P16457 cell
    division protein ftsW (414 aa), fasta scores: opt: 647
    z-score: 573.9 E(): 7.7e-25, 37.0% identity in 370 aa
    overlap. Contains PS00428 Cell cycle proteins ftsW / rodA
    / spoVE signature and PS00041 Bacterial regulatory
    proteins arac family signature"
    /codon_start=1
    /transl_table=11
    /product="cell division protein ftsW"
    /protein_id="CA18669.1"
    /db_xref="gi:3080474"
    /db_xref="sptrembl:069553"
    /translation="MGNALGRDRGRTGNNNSVOAATVEAVEGRTATGSEVSEAK
    RGSIRACVRRPOTRPDAPGRPMASPHILIAVALLRLGILVLSGSRISGDG
    NSMTIFGQVYAGMTTIGICGTSSLMHSIFIRIRAFSTVITLILVLPVIGINLA
    LVAQPDIGQVVSJGJITLLALVAGLPLRVITISLAVFAGAILMSAGYSERAR
    SMNAPEDPODTGQVQARQKALAHGIGDVGQVAKMYLPLNANDEIFAIIEE
    LGLVAGLPLIGLFGITCMRLARRSADPFLRLTATTVMVVGQAFINIGYIGVL
    PYTGOLPFIISAGTSAAILFMIGIMANAARHEPAVALRAGDKVNOILRLPLP
    VYLLPVRVETRDROKVRPOPGKASVGRPPARKARPRDRPLRAPFVADRLRHNSG
    QRRAGOLRAGQRTYPRVPRVEGGRYC"
    /complement(1937..1942)
    /gene="ftsW"
    /note="possible RBS upstream of murG"
    2019..2020
    /note="conflict: AG is AAG in MUB27CS"
    2053
    /note="conflict: this base is missing in MUB27CS"
    2061..2062
    /note="conflict: TC is CA in MUB27CS"
    2081
    /note="conflict: this base is missing in MUB27CS"
    2163
    /note="conflict: G is T in MUB27CS"
    /complement(2300..2428)
    /gene="ftsW"
    /note="PS00041 Bacterial regulatory proteins, arac family
    signature"
    /complement(2303..2377)
    /gene="ftsW"
    /note="PS00428 Cell cycle proteins ftsW / rodA / spoVE
    signature"
    2312..2313
    /note="conflict: AC is AAC in MUB27CS"
    2973..2974
    /note="conflict: AC is AAC in MUB27CS"
    2989
    /note="conflict: this base is missing in MUB27CS"
    /complement(3529..3534)
    /note="possible RBS upstream of ftsW"
    /complement(3533..5005)
    /gene="murD"
    /complement(3533..5005)
    /gene="murD"
    /note="MUB268.04c, murD, probable

```

Query Match 14.3%; Score 226.2; DB 1; Length 38859;  
 Best Local Similarity 58.3%; Pred. No. 2.8e-44;  
 Matches 396; Conservative 0; Mismatches 283; Indels 0; Gaps 0;

udp-n-acetylmutramoylalanine--d-glutamate ligase, len: 490

QY 137 gatttgcgcgcgtccatctatggtatttggcggagccggatgtcgttgcgcga 196  
 DB 681 GAGTTGACGCGGTGCACATGTGGGATTCGGGATTCGGGATTCGGGATTCGGG 622  
 QY 197 attctgtctcgcgcgttaagacagctgttccgattcccaagaatcccgaccttg 256  
 DB 621 ATCTGTTAGACGCGGGCGGTGTGGTGTCTGCTCAATTCGCAAGAGTCGGTGTG 562  
 QY 257 ctctcactcgcgcgcgttggagccacatcgcagtggaacgcgtcggaaaaccttga 316  
 DB 561 CACGACTGCGGGCGCGCGCGCTGATCCGATCCGATCCGATCCGATCCGATCCG 502  
 QY 317 ctctcgcgcgcgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 376  
 DB 501 TTGCTGCGTGGCGCGCTACCGAGTACATCCGATCCGATCCGATCCGATCCG 442  
 QY 377 ccggaacttcttgcgcgcgttgaagaagcgttgcgttgcgttgcgttgcgttgc 436  
 DB 441 CCGGCTGCTGCGAGCGCGAGCGCGCGCTGATCCGATCCGATCCGATCCGATCC 382  
 QY 437 ggcgaattgtcgttgcgcgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 496  
 DB 381 GCCAGCTGATGATGGGTGTACACAGCTGATGCTGTGCTGACACACGCAAGACTCG 322  
 QY 497 acccctctatgtctgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 556  
 DB 321 ACAGCTGATGCTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 262  
 QY 557 ggcgaagatctcaacaaagcttgcgcgcgttgcgcgcgttgcgcgcgttgcgcgc 616  
 DB 261 GCGGCTGATGCTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 202  
 QY 617 gctaaagcagatgaatctacgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 676  
 DB 201 GCCAGCTGATGCTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 142  
 QY 677 accaattgtgaacacagacacgttgcgttgcgttgcgttgcgttgcgttgcgttgc 736  
 DB 141 ACCAAGCTGATGCTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 82  
 QY 737 ttcaagatttgcgcgcgttgcgcgcgttgcgcgcgttgcgcgcgttgcgcgcgttgc 796  
 DB 81 TTCAAGCTGCTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 22  
 QY 797 cctcaagcagcgcgcgttgcgcgcgttgcgcgcgttgcgcgcgttgcgcgcgttgc 815  
 DB 21 CCGGCTGATGCTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3

RESULT 15  
 AF002997 329709 bp DNA BCT 15-MAY-2001  
 LOCUS Mesorhizobium loti DNA, complete genome, section 4/21.  
 DEFINITION AF002997 BA000012  
 ACCESSION AF002997.2 GI:14022051  
 VERSION  
 KEYWORDS  
 SOURCE Mesorhizobium loti (strain:MAFF303099) DNA.  
 ORGANISM Mesorhizobium loti  
 Bacteria: Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 Phyllobacteriaceae; Mesorhizobium.  
 REFERENCE  
 AUTHORS Kaneko, T., Nakamura, Y., Sato, S., Asamizu, E., Kato, T., Sasamoto, S.,  
 Watanabe, A., Idesawa, K., Ishikawa, A., Kawashima, K., Kimura, T.,  
 Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A.,  
 Mochizuki, T., Nakayama, S., Nakazaki, N., Shimo, S., Sugimoto, M.,  
 Takeuchi, C., Yamada, M. and Tabata, S.  
 TITLE Complete genome structure of the nitrogen-fixing symbiotic

JOURNAL bacterium Mesorhizobium loti  
 MEDLINE DNA Res. 7 (6), 331-338 (2000)  
 REFERENCE 21082930  
 AUTHORS Kaneko, T.  
 TITLE Direct Submission  
 JOURNAL Submitted (05-DEC-2000) Takakazu Kaneko, Kazusa DNA Research  
 Institute, The First Laboratory for Plant Gene Research, Yana  
 153-3, Kisarazu, Chiba 292-0812, Japan  
 (E-mail:kaneko@kazusa.or.jp,  
 URL:http://www.kazusa.or.jp/rhizobase/  
 Tel:81-438-52-3935(fax.2338) Fax:81-438-52-3934  
 On May 11, 2001 this sequence version replaced gi:11994965.  
 COMMENT  
 FEATURES  
 SOURCE

1. 329709  
 /organism="Mesorhizobium loti"  
 /strain="MAFF303099"  
 /db\_xref="taxon:381"  
 complement(79..402)  
 /gene="ml11244"  
 complement(79..402)  
 /note="unknown protein"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="BAB48662.1"  
 /db\_xref="GI:14022052"  
 /translation="MSRGLSIPFNGSFHWADCLAVORPINSDLLSIFPAFMVR  
 LLKGPWRNRPQYATILGAVVAMLLNGLMPAYDDPFVIGVTGIFGSMAMALFDA  
 ILGTA"  
 411..2480  
 /gene="ml1245"  
 411..2480  
 /gene="ml1245"  
 /codon\_start=1  
 /transl\_table=11  
 /product="probable electron transfer protein"  
 /protein\_id="BAB48663.1"  
 /db\_xref="GI:14022053"  
 /translation="MNSPANTIDPLVLEPMSGKRGKRPVGTPIVLDARGLGYVESVC  
 GGRATGCOIEVQENFPAKRTIVSSNDHISPKGKRETEREYRGLEPERRLSGSAOI  
 LGDVIDPQDVIINAOTIRKADTRVIAKOTAIKMCYVEIEEPDMKPLGSLDLKI  
 ALAMDGFKNLEDFYLLPOVGGILRKGNMTATPAHKDADSDIAVIALMPGLRNEA  
 YGLACDIDGSTTIAMHLVSLISGRVNASGTSNPQIRGEDLMRSVYVMANPDGREGM  
 TVAVREAISSLDVCKACBEGNVORNDIISVGVNPIMHNLFLIDPTELGAPFALAV  
 SGAVRIKASDIDGLKNGARLEYLPCIGHVGAADAAVATIESEGHRODEMMLIVGT  
 NAEIVLGNRRARVVAASPTGPAFEGAEIISGGORAPPAIERVVIDDPTLEPKYRYS  
 ELMSEDEGFLDSVQATCYTCIGSSIIETIVEMTLAIISEDEVYSGSLAASPRVTA  
 NGRTFVYLNKGEPRKITTITDVRATIQAKALATGKLLMERKONTEDVRIHFGAF  
 GSFIDPKYAEVGLIPDCDLNVSAVGAAGARMLLNKRGIRREIETVSOIKIE  
 TALEPKQEHFVYAMALPNKVDPPFKLSAAVKKPRTVSEDIAGDAPRRRSREH  
 AARSRE"  
 2529..2852  
 /gene="ml1246"  
 2529..2852  
 /gene="ml1246"  
 /note="unknown protein"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="BAB48664.1"  
 /db\_xref="GI:14022054"  
 /translation="MFPDRTPDETETLFSRNLSDTQWAPRIETRPETIILGRSRE  
 MKMSFENLGRFGVTIRONHAKRAIRALNSLPAHVQKIDIGWAPSPDQAAIASL  
 LGTAR"  
 3087..4040  
 /gene="ml1247"  
 3087..4040  
 /gene="ml1247"  
 /codon\_start=1  
 /transl\_table=11  
 /product="esterase"  
 /protein\_id="BAB48665.1"  
 /db\_xref="GI:14022055"



```

Db 248902 CATCATGCTATGGACCAATGGCGATGGGAGAGGAGTGGATGGTGTGAGGC 248843
QY 625 agatgaatctgaacgcatgctgctgcttaagcaaatgtgcagtgttacaacatgt 684
Db 248842 CGACGAGAGCGACGCACTTCTCAAAATTCGCGGACATGCGCGTGTCAACACAT 248783
QY 685 ggaacagaccacttgacttcttaaaacccctgaagccctacttcaagtgttcagca 744
Db 248782 CGATCCGACATCTGACCACTATGGCAGTTTCAGACAAAGTGGCGAGGCGTTCGCGCA 248723
QY 745 ttgcagagagcatcaaccggaagcaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 804
Db 248722 GTTCGTGAGAACGTGCGCTTCTACGCTTTGGCGTGTGTGACCCGACCATCTGAGAT 248663
QY 805 agcggagctggyggaagaggtctgtccgcaagygatcaagaactgtgtgtgtgtgtgt 864
Db 248662 GCAGGCGCTGTGTGCGCGCATGAGAGACGCGCGGTATCATCTACGCGAGAACGCGCA 248603
QY 865 tgaagccgttacagcaaccctgaggttccagagatgctacatgctgtgtgtgtgtgtgt 924
Db 248602 AGCCGATGTGCGCTTACCAACCAACCGCATGGCGGCGCCACC-----TCCGAAAT 248552
QY 925 tgtgcagaagagcaccgcyccacatacatgatgtgacaggaagatctgtgtatct 984
Db 248551 CGACGTGTGATCCGGACCGCAAGACCGGACGACAGCAGATTTAGGCTGTGCGCT 248492
QY 985 tcaatccctggtgcatcacatgttactaacggtgcagccctgtgtgtgtgtgtgtgtgt 1044
Db 248491 GCCGATGCCGCGCGCGCAATGTCTCCAAAGCCACCGCGCATGCGGTGCGCATGA 248432
QY 1045 ggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1104
Db 248431 GCTGCGCTGTGCGGCGAGCATCAAGAGGCGCTGTGCTTGTGCGGCGGTCAAGCG 248372
QY 1105 ccgcttctgagttccacggtgtctatcgagggcgcaaatctaattgagcgtctattatga 1164
Db 248371 CCGCTT-----CACCCATACCGGATCTTGGAAGGCGGTGATGTGTGA 248327
QY 1165 tgattacgacacaccacaaggaagttaactgacgtgtgtgtgtgtgtgtgtgtgtgtgt 1224
Db 248326 CGATTACGCCACCATCTCGGTGAGATCTCGCGGTGTGTGAGGCGCGCGCA----- 248272
QY 1225 gaagccgctggaagagcggtgtcatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1284
Db 248273 -----GCGCACCAAGAGGCGCGGTATCGCATTTGCCAACCGCACCGCTTACCCGGCT 248219
QY 1285 catagaatttcaaaagaggttgcggygggcaactgtacgtgcagagcgtgcgtgtgtgt 1344
Db 248218 CCACGATCTGTTCAGAGATTTCTCCGCTGTCTTCAACGACGCCGACACGATGTGTGC 248159
QY 1345 tgaatttcaagagcgcgcaacaacggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1392
Db 248158 GCCGCTATGTGTGCGGAGAACCGATCGACGGGTGAGCTCGGA 248111

```

Search completed: February 15, 2002, 08:59:52  
 Job time: 4502 sec